STIC-ILL

From: Ramirez, Delia
Sent: Thursday, March 18, 2004 7:17 PM
To: STIC-ILL
Subject: case 09/966742

Hi,

I would like to request photocopies of the following material:

Author: Duerfahrt T, Doekel S, Sonke T, Quaedflieg PJ, Marahiel MA.
Title: Construction of hybrid peptide synthetases for the production of alpha-l-aspartyl-l-phenylalanine, a precursor for the high-intensity sweetener aspartame.
Journal: Eur J Biochem. 2003 Nov;270(22):4555-63.

Author: Doekel S, Eppelmann K, Marahiel MA.
Title: Heterologous expression of nonribosomal peptide synthetases in B. subtilis: construction of a bi-functional B subtilis/E coli shuttle vector system.
Journal: FEMS Microbiol Lett. 2002 Nov 5;216(2):185-91.

Author: Doekel S, Marahiel MA.
Ttile: Biosynthesis of natural products on modular peptide synthetases.
Journal: Metab Eng. 2001 Jan;3(1):64-77. Review.

Author: Doekel S, Marahiel MA.
Title: Dipeptide formation on engineered hybrid peptide synthetases.
Journal: Chem Biol. 2000 Jun;7(6):373-84

Thank you,

Delia M. Ramirez, Ph.D.
Patent Examiner
Recombinant Enzymes-Art Unit 1652
USPTO
400 Dulany Street, Remsen Bldg., 3A74, Mail room 3C70
Alexandria, VA 22314
(571) 272-0938
delia.ramirez@uspto.gov

ADONIS - Electronic Journal Services

Requested by

Adonis

Article title	Biosynthesis of natural products on modular peptide synthetases
Article identifier	1096717601000337
Authors	Doekel_S Marahiel_M_A
Journal title	Metabolic Engineering
ISSN	1096-7176
Publisher	Academic Press USA
Year of publication	2001
Volume	3
Issue	1
Supplement	0
Page range	64-77
Number of pages	14
User name	Adonis
Cost centre	
PCC	$35.00
Date and time	Friday, March 19, 2004 7:52:23 PM

Copyright © 1991-1999 ADONIS and/or licensors.

The use of this system and its contents is restricted to the terms and conditions laid down in the Journal Delivery and User Agreement. Whilst the information contained on each CD-ROM has been obtained from sources believed to be reliable, no liability shall attach to ADONIS or the publisher in respect of any of its contents or in respect of any use of the system.

18069302

Metabolic Engineering 3, 64–77 (2001)
doi:10.1006/mben.2000.0170, available online at http://www.idealibrary.com on IDEAL®

REVIEW

Biosynthesis of Natural Products on Modular Peptide Synthetases

Sascha Doekel[1] and Mohamed A. Marahiel[2]

Fachbereich Chemie/Biochemie, Philipps-Universität Marburg, Hans-Meerwein-Straße, 35032 Marburg, Germany

Received January 5, 2000; accepted October 3, 2000; published online December 29, 2000

Microbial nonribosomally processed peptides represent a large class of natural products including numerous important pharmaceutical agents, as well as other representatives that play a prevalent role in pathogenicity of certain microorganisms [M. A. Marahiel, T. Stachelhaus, and H. D. Mootz (1997). *Chem. Rev.* 97, 2651–2673]. Although diverse in structure, nonribosomally synthesized peptides have a common mode of biosynthesis. They are assembled on very large protein templates called peptide synthetases that exhibit a modular organization, allowing polymerization of monomers in an assembly-line-like mechanism. © 2001 Academic Press

INTRODUCTION

Nonribosomal peptide synthetases (NRPSs) are organized in iterative functional units called modules that catalyze a handful of catalytic reactions leading to peptide formation. A minimal elongation module that carries out one elongation step comprises at least three catalytic domains: one for substrate recognition and activation by adenylation [adenylation domain (Conti *et al.*, 1997)], one for covalent incorporation as thioesters on an enzyme-bound 4′-phosphopantheinyl (4′PP) moiety [thiolation domain (Stachelhaus *et al.*, 1996)], and one for condensation of the precursor (condensation domain) leading to formation of a peptide concatenated to the 4′PP moiety of the next module (Stachelhaus *et al.*, 1998). Insertions of optional catalytic domains into a minimal module may result in further substrate modification including epimerization of α-carbon atoms [epimerization domain (Stindl and Keller, 1994)] or N-methylation of α-amino groups [N-methylation domain (de Crécy-Lagard *et al.*, 1997)]. Condensation of two amino acid precursors can also lead to heterocyclic ring formation of oxazoline or thiazoline types (Gehring *et al.*, 1998b). Such reactions involving side chains of substrate amino acids like threonine, serine, or cysteine are accomplished by inherent peptide synthetase domains called cyclization domains (Konz *et al.*, 1997). Release of the processed peptide from the last module's 4′PP moiety is implemented by thioesterase-like domains resembling analogous domains from fatty acid synthases (Shaw-Reid *et al.*, 1999). Nonribosomal synthesized peptides are released from the template as linear, cyclic, or branched cyclic molecules. Release of the thioester-bound peptides can also be catalyzed by NADPH-dependent reduction to peptidyl amino aldehydes (Ehmann *et al.*, 1999) or peptidyl amino alcohols (Billman-Jacobe *et al.*, 1999). Taken into account that nonribosomal peptide synthesis is not restricted to proteinogenic amino acids—a few hundred carboxy acids are known to be incorporated—together with the set of possible modifications during assembly an immense number of peptides and peptide analogous can be designed by nature employing a single group of proteins (Konz and Marahiel, 1999).

Due to the modular arrangement of peptide synthetase these biofactories have an enormous size (Weber *et al.*, 1994)—the largest known peptide synthetase directing the synthesis of the immunosuppressive agent cyclosporin contains 11 modules on a single polypeptide chain and has a mass of 1600 kDa. An illustration of the stepwise N- to C-terminal assembly of nonribosomal peptides is given in Fig. 1, depicting the biosynthesis of bacitracin in *Bacillus licheniformis* (Konz *et al.*, 1997). In most cases, primary sequence of the processed peptide and arrangement of the modules display a colinearity underlining the assembly-line mechanism (see an exception from this rule in biosynthesis of syringomycin below).

Diversity of nonribosomally processed peptides is not restricted to their structure; their biological impact is also impressive. Important pharmacological agents like the immunosuppressive cyclosporin (Weber *et al.*, 1994) and the peptide backbones of antibiotics of the vancomycin (van Wageningen *et al.*, 1998), penicillin, and cephalosporin (Gutierrez *et al.*, 1991; Smith *et al.*, 1990) group are produced by peptide synthetases (see Table 1). Some

[1] E-mail: doekel@chemie.uni-marburg.de.
[2] To whom correspondence and reprint requests should be addressed. Fax: 49-6421-2822191. E-mail: marahiel@chemie.uni-marburg.de.

1096-7176/01 $35.00
Copyright © 2001 by Academic Press
All rights of reproduction in any form reserved.

Metabolic Engineering 3, 64–77 (2001)
doi:10.1006/mben.2000.0170

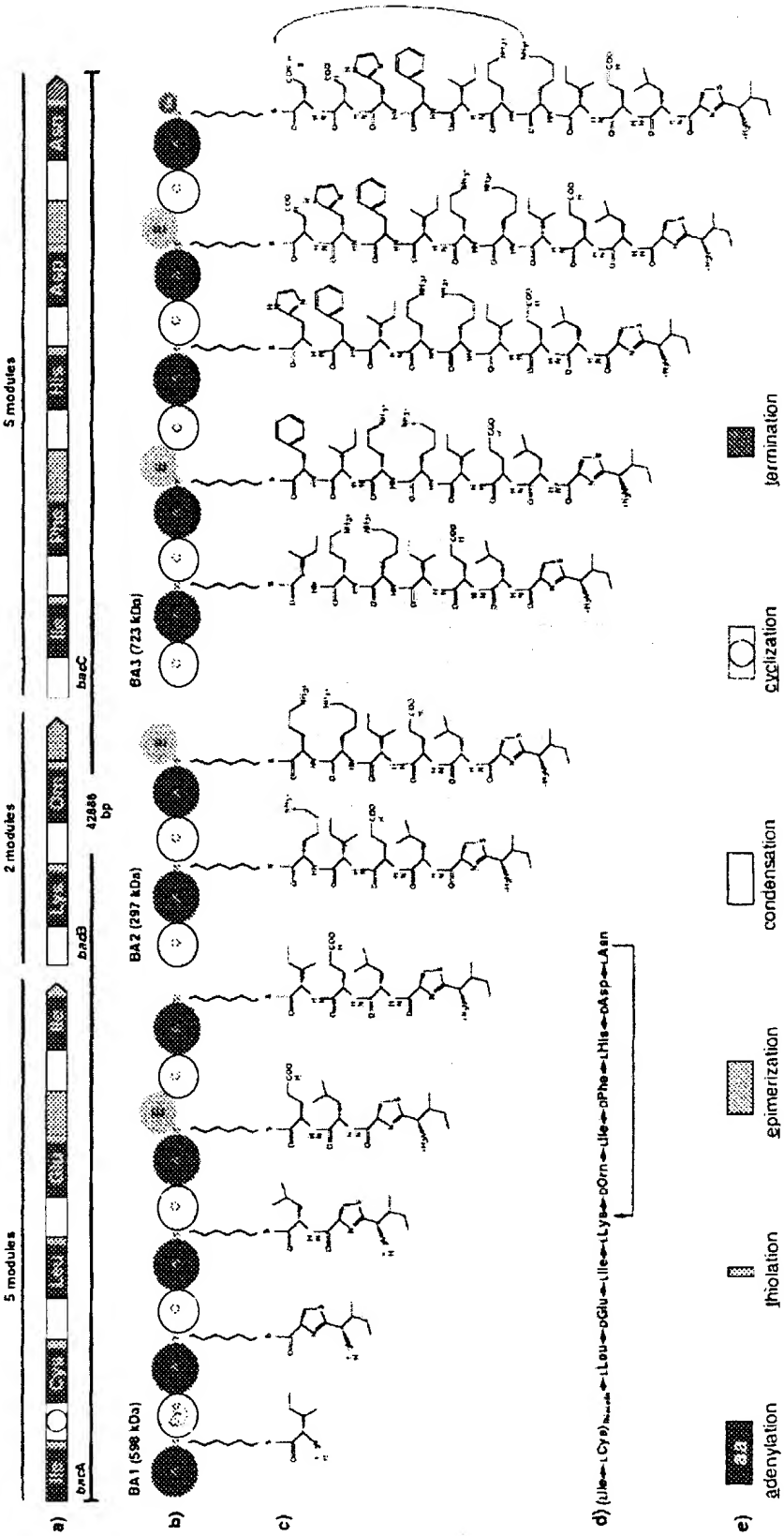


FIG. 1. The multiple carrier thiotemplate model illustrated by the biosynthesis of bacitracin. (a) Genes *bacA*, *bacB*, and *bacC* that are part of an operon. The domain structure of the corresponding proteins is indicated by different shadings. (b) Illustration of the domain structure of BA1, BA2, and BA3. Note the repetitive character of the proteins: BA1 and BA3 contain five modules each, while BA2 contains two modules. Recognition and activation of amino acid substrates via their adenylates are accomplished on the adenylation domains before the activated amino acids become concatenated to the cofactor 4′-phosphopantetheine which is part of each module's thiolation domain. (c) The stepwise N- to C-terminal assembly of the peptide is shown. The reactions are performed by the condensation domains or, alternatively, by the cyclization domains forming thiazoline heterocycles. Intermediate epimerization domains convert appropriate amino acids into D-configuration (positions 4, 7, 9, and 11). The peptide is cleaved from the enzyme by the action of a C-terminal thioesterase-like domain. (d) Primary sequence of the processed peptide.

Metabolic Engineering 3, 64–77 (2001)
doi:10.1006/mben.2000.0170

peptides play a role in pathogenicity of certain microorganisms: the hepatotoxins of the cyanobacterium *Microcystis aeruginosa* (Dittmann *et al.*, 1997) as well as several fungal toxins are produced by peptide synthetases (see Table 1) (Scott-Craig *et al.*, 1992; Tudzynski *et al.*, 1999). Also, the production of ferric-chelating siderophores like vibriobactin (Wyckoff *et al.*, 1997) by enteric microorganisms depends on peptide synthetases. Their products therefore play a prevalent role in pathogenicity of the organisms causing wound infections, plague, cholerae, or tuberculosis. Lipopeptides like surfactin of *Bacillus subtilis* have also awoken interest because of their biotechnological relevance in oil-recovery processes (Fiechter, 1992).

Recently a large number of genes encoding peptide synthetases have been identified. An approach taking advantage from the ubiquitous presence of highly conserved core sequences has been devised to identify unknown peptide synthetase genes in several microorganisms (Turgay and Marahiel, 1994). Also, genome-sequencing projects of model organisms provide us with an increasing number of peptide synthetase genes.

PEPTIDE SYNTHETASES ARE COMPOSED OF CATALYTIC DOMAINS

Sequence comparisons within NRPSs helped to define catalytic domains within each module. It could be shown that these domains can be considered as semiautonomous units that retain their potential of catalyzing part reactions of nonribosomal peptide synthesis when overproduced as distinct proteins (Mootz and Marahiel, 1997; Stachelhaus *et al.*, 1996).

Compared to the ribosomal protein machinery, NRPSs are less specific. Probably the evolutionary force to maintain a given substrate selection is not very strict. The main gatekeeper for specificity of NRPSs is the adenylation domain that catalyzes the activation of cognate carboxy acids as adenylates by hydrolysis of ATP. Adenylation domains that are ca. 550 amino acids in size represent therefore a member of the superfamily of adenylate-forming enzymes like firefly luciferases and acyl-CoA-ligases. Usually adenylation domains in NRPSs are found as part of a larger protein frame but it could be demonstrated that they exhibit a comparable catalytic activity also when expressed heterologously as separated units (Mootz and Marahiel, 1997). The importance of highly conserved residues within the A-domains for catalytic efficiency was demonstrated by site-directed mutagenesis (Gocht and Marahiel, 1994) and photoaffinity labeling experiments (Pavela-Vrancic *et al.*, 1994). The most impressive insight into the architecture of adenylation domains can be drawn from the crystal structure of the first adenylation domain of the gramicidin S synthetase GrsA (Conti *et al.*, 1997). Although a low homology exists in primary structure of NRPS adenylation domains and firefly luciferases, their overall three-dimensional structure is almost identical (Denessiouk and Johnson, 2000). Because adenylation domains of NRPSs share a sequence identity of about 30–60%, the structure of GrsA can be anticipated as prototype for all adenylation domains. The crystal structure underlines the importance of nine residues in substrate recognition. A detailed alignment of these residues in other adenylation domains with diverse specificity allowed establishment of a nonribosomal code (Stachelhaus *et al.*, 1999). Altering the specificity of adenylation domains in a predictable way opens new directions to change the structure of natural compounds.

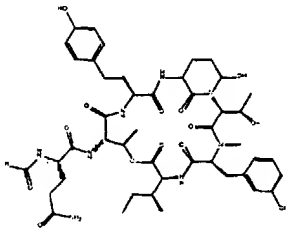
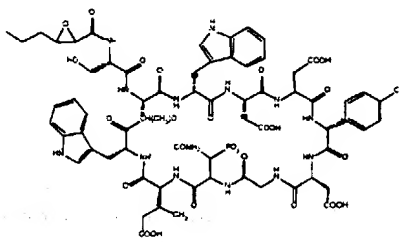
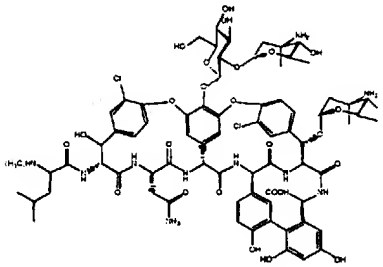
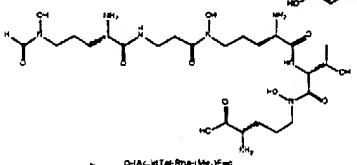
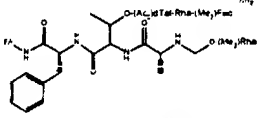
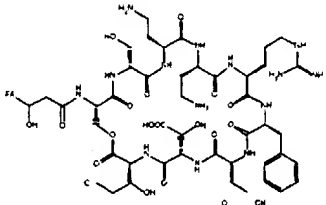
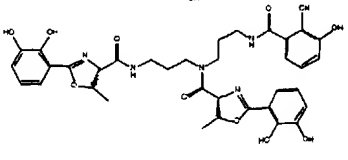
A small protein domain of ca. 80 residues that follows the adenylation domain is called the thiolation domain (Stachelhaus *et al.*, 1996). This structural-independent unit displays a highly conserved serine residue that serves as the attachment site for the cofactor 4′-phosphopantetheine. The thiol moiety of the cofactor is the carrier for the previously activated carboxy acid. In analogy to fatty acid and polyketide biosynthesis, this domain is also designated as peptidyl carrier protein (PCP). Since it has been elucidated that each module contains a PCP domain and therefore a 4′-phosphopantetheine arm, the initial model of NRPS (which suggested only a single, central cofactor) has been rendered to the now accepted multiple carrier model (Stein *et al.*, 1996). The recently determined solution structure of a PCP domain from the tyrocidine synthetase module TycC3 revealed a common fold for ACP and PCP domains (Weber *et al.*, 2000).

The presence of the cofactor 4′-phosphopantetheine—the state of a holo-enzyme—is crucial to nonribosomal peptide synthesis (Nakano *et al.*, 1992). This specific phosphopantetheinylation is implemented by CoA-binding enzymes of the superfamily of 4′PP transferases (Lambalot *et al.*, 1996). The enzymatic reaction results in phosphate ester formation of the hydroxyl group of the invariant serine residue of ACP and PCP domains with the 4′-phosphopantetheinyl moiety. Genes encoding for 4′PP transferases in prokaryotes are often found adjacent to peptide synthetase operons; nevertheless, when an organism contains two or more NRPS systems it seems likely that one 4′PP transferase is sufficient to modify both of them. 4′PP transferases have been the subject of extended biochemical and structural investigations: PCP and ACP domains can be transferred to holo-form by 4′PP transferases in the presence of CoA and magnesium *in vitro* (Weinreb *et al.*, 1998). Also, posttranslational modification of NRPS to holo-enzyme *in vivo* was shown when coexpression with genes encoding 4′PP

Review

Metabolic Engineering 3, 64–77 (2001)
doi:10.1006/mben.2000.0170

TABLE 1

Product	Organism	Gene	Organization	Accession	Literature	Primary Structure of Product
Bacterial peptide synthetase clusters:						
Anabaenopeptolide	*Anabaena strain90*	apdA apdB apdD	FAT CAT CAT CAT CAMT CAMT CAT Te		[Rouhiainen, 2000 #1098]	 1)
CDA†	*Streptomyces coelicolor A3(2)*	SCE 63.03c SCE 63.02c SCE 63.01c	CAT CAT CATE CAT CAT CATE (Ser Thr Trp Asn Asp HPG) CAT CAT CATE (Asp Gly (3-P)Asn) CAT CAT Te (3-MeGlu Trp)	AL035640.2	[Chong, 1998 #11; Redenbach, 1996 #354]	
Chloroerenomycin†	*Amycolatopsis orientalis*	PCZA 363.3 PCZA 363.4 PCZA 363.5	AT CATE CAT (DLeu Tyr Asn) CATE CATE CAT (HPG HPG Tyr) CAT C Te (DHPG)	AJ223999.1	[van Wageningen, 1998 #353; Trauger, 2000 #1086]	
Exochelin§	*Mycobacterium smegmatis*	*fxb*B *fxb*C	AT CATC (Orn *beta*Ala) ATE°CATE CAT°CAT Te (Orn Thr Orn ?)	AF027770.1	[Fiss, 1994 #355; Yu, 1998 #342; Zhu, 1998 #343]	
Glycolipopeptid	*Mycobacterium smegmatis*	*mps*	CATE CATE CATE CAT R (Phe *allo*Thr Ala Ala?)	AJ238027.1	[Billman-Jacobe, 1999 #357]	
Syringomycin†	*Pseudomonas syringae*	*syr*B *syr*E	AT (Thr) CAT CAT CAT CAT CAT CAT CAT (Ser dSer Dab dDab Arg Phe Dh-Thr) CAT CT Te (Asp)	U25130.2 AF047828.1	[Guenzi, 1998 #345]	
Vibriobactin§	*Vibrio anguillarum*	*vib*E *vib*B *vib*F *vib*H	A (DHB) Isochorismatase+T Z ZAMT (C) (Thr) C	O07900 AF030977.2	[Wyckoff, 1997 #358; Butterton, 2000 #1093]	

Metabolic Engineering 3, 64–77 (2001)
doi:10.1006/mben.2000.0170

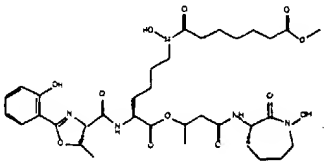
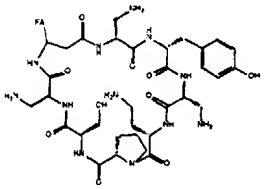
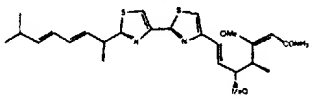
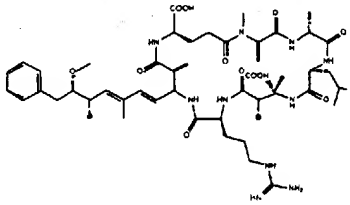
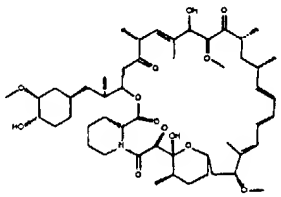
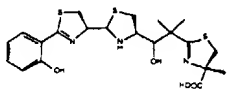
Fungal peptide synthetase clusters:

ACV[1]	*Penicillium chrysogenum*	*pch*AB	AT CAT CATE Te Aad Cys Val	P19787	[Smith, 1990 #1051; Schwecke, 1992 #179; Gutierrez, 1991 #903]
AM-Toxin[2]	*Alternaria alternata*	*AMT*	CAT CAT CAT CAT C ? ? ? ?	AF184074.1	[Higashijima, 1979 #1104; Johnson, 2000 #1075]
Cyclosporin[3]	*Tolypocladium inflatum*	*css*A	AT CAMT CAMT CAMT CAMT CAT OAla Leu Leu Val Bmt Abu CAMT CAMT CAT CAMT CAT Gly Leu Val Leu Ala	S41309	[Weber, 1994 #360]
Enniatin[4]	*Fusarium scirpi*	*esyn*1	AT CAMT TC OHiV Val	S39842	[Haese, 1993 #365; Haese, 1994 #364; Hacker, 2000 #1074]
Ergotamin[5]	*Claviceps purpurea*	*ps*1	AT CAT CAT C Ala Phe Pro	AJ01196.4	[Tudzynski, 1999 #340; Riederer, 1996 #339]
Ferrichrome[6,§]	*Ustilago maydis*	*sid*2	AT CAT CAT CT C ? ? ?	O43103	[Mei, 1993 #410]
HC-toxin[8]	*Cochliobulus carbonum*	*hts*1	ATE CAT CAT CAT Pro Ala Ala Aeo	Q018886	[Scott-Craig, 1992 #366]
Lysinbiosynthese	*Candida albicans*	*lys*2	AT R Aad	Q12572	[Suvarna, 1998 #368; Ehmann, 1999 #303]

Mixed NRPS/PKS clusters:

Epothilon[9]	*Sorangium cellulosum So c90*	epoA	KS AT ER ACP	AF210843	[Tang, 2000 #1102; Julien, 2000 #1103; Molnar, 2000 #955]
		epoP	ZA Oxy T Cys		
		epo B	KS AT DH KR ACP ?		
		epoC	KS AT KR ACP KS AT KS AT DH ER KR ACP KS AT DH ER KR ACP		
		epoD	KS AT KR ACP KS AT DH MT KR ACP		
		epoE	KS AT DH ER KR ACP Te		
FK506[3]	*Streptomyces sp.*MA6548	*fkb*P	CAT C Pip	AF082100.1	[Motamedi, 1998 #379]

Metabolic Engineering 3, 64–77 (2001)
doi:10.1006/mben.2000.0170

Mycobactin§	*Mycobacterium tuberculosis*	*mbtA*	A (Sal)	Z95208.1	[Cole, 1998 #369; Quadri, 1998 #44]	
		mbtB	T ZAT AT (Thr)			
		mbtC	KS			
		mbtD	ACP AT KR ACP			
		mbtE	CAT CT (Lys)			
		mbtF	CATE (Lys)			
Mycosubtilin†	*Bacillus subtilis* ATCC 6633	*mycA*	CoA-S ACP KS ACP AaT CT CAT C (Asn)	AF184956.1	[Duitman, 1999 #338]	
		mycB	(C)ATE CATE CAT CAT C (Tyr Asn Gln Pro)			
		mycC	(C)ATE CAT Te (Ser Asn)			
Myxothiazol†	Stigmatella aurantiaca DW4/3	mtaB	ACP KS AT AT DH ER KR ACP KS AT DH KR ACP	AF 188287.1	[Beyer, 1999 #412; Silakowski, 1999 #413]	
			ZAT Oxy (?)			
			ZAMT KS AT DH KR ACP (?)			
			KS AT MT KR ACP			
			KS AT MT ACP			
			CA Oxy T Te			
Microcystin *	*Microcystis aeruginosa*	*mcyA*	AMT CATE (MdhA Ala)	AB019578.1/ AF183408.1	[Neilan, 1997 #370; Neilan, 1999 #3; Nishizawa, 1999 #327]	
		mcyB	CAT CAT^{D} (Leu 3-MeAsp)			
		mcyC	CAT Te (Arg)			
		mcyD	KS AT DH MT KR ACP KS AT KR ACP			
		mcyE	KS AT MT AaT T CAT C (?)			
		mcyG	AT KS AT MT KR (?)			
Rapamycin^	*Streptomyces hydroscopius*	*rapP*	CAT C (Pip)	X86780.1	[Schwecke, 1995 #371]	
Yersiniabactin§	*Yersinia pestis*	*ybtS*	A (Sal)	U50364.1 AF09125.1	[Bearden, 1997 #376; Gehring, 1998 #46]	
		irp2	TZAMT ZT (Cys)			
		irp1	KS AT MT KR ACP ZMT Te			

Legend:
† antibiotic; § siderophore; Δ immunosuppressant; * phytotoxin; # cyctostatic agent. Domains of peptide synthetases in bold: **A**, adenylation domain; **T**, thiolation domain; **C**, condensation domain; **E**, epimerization domain; **Z**, heterocyclization domain; **M**, N-methyltransferase domain; **Te**, thioesterase domain; **R**, reductase domain; **F**, N-formylase domain. Domains of polyketide synthases: **KS**, ketosynthase; **AT**, acyltransferase; **MT**, methyltransferase; **KR**, ketoreductase; **DH**, dehydrogenase; **ACP**, acylcarrierprotein; **AaT**, amino acid transferase; brackets indicate truncated domains. Substrate amino acids are abbreviated with three-letter-code, modified amino acids are indicated by the use of prefixes: D, D-configuration; allo, allo-configuration; Me, N-methylation; xP, O-phosporylation at position x related to the α carbon atom; h, dehydro; xh, hydroxylation at position x related to the α carbon atom; xf, formylation at position x related to the α carbon atom. Greek letters indicate position of peptide bond related to the α carbon atom. Abbreviations of non-proteinogenic amino acids and hydroxy acids: Aad, amino-adipic acid; Aeo, 2-amino-8-oxo-9,10-epoxy decanoic acid; Bmt, 4-(2-butenyl)-4-methyl-threonine, Dab, 2,4-diamino-butyric-acid; DHB, dihydroxy benzoate; HIV, 2-hydroxy-isovalerat; HPG, 4-hydroxy-phenylglycine; Mdha, N-methyldehydro-alanine; Orn, ornithine; Pip, pipecolic acid; Sal, salicylate. FA, fatty acid. [1] Anabaenopeptolide. **B** [2] structure from *Alternaria mali* according to [Higashijima, 1979 #1104].

Metabolic Engineering 3, 64–77 (2001)
doi:10.1006/mben.2000.0170

transferases was performed (Kealey *et al.*, 1998). Certain 4'PP transferases do not exhibit a restrictive substrate specificity toward the PCP or ACP domain so that heterologous thiolation domains can be modified to holo-forms (Lambalot *et al.*, 1996). The recently determined crystal structure of the 4'PP transferases Sfp (28 kDa) of surfactin synthetases in *B. subtilis* revealed an unexpected intramolecular twofold pseudosymmetry of the protein (Reuter *et al.*, 1999). This structure may provide an explanation for the observation that the homologous 14-kDa ACPS (ACP-synthase of *Escherichia coli*) may function as homodimer.

The stepwise N- to C-terminal condensation of the thioesterified carboxy acids is accomplished by a domain located N-terminal of adenylation domains. The role of this condensation domain in the formation of a peptide bond has been clearly assigned in an *in vitro* study using a truncated dimodular model system (Stachelhaus *et al.*, 1998). This domain type is ca. 450 amino acids in length. In general, the number of condensation domains within a peptide synthetase is consistent with the number of peptide bonds that are formed (see below). In contrast to adenylation and thiolation domains, condensation domains display a protein class that has no extended sequence homologies to proteins of known mechanisms. The presence of a highly conserved motif HHxxDGxS (His motif) that is also found in acyl transferases like chloramphenicol acetyltransferases and lipoyl transferases suggests a mechanism in which one of the histidines may enhance the nucleophilic character of the amine nitrogen to facilitate formation of the peptide bond (De Crécy-Lagard *et al.*, 1995). The importance of the second histidine in catalysis of peptide bond formation could be shown by site-directed mutagenesis (Stachelhaus *et al.*, 1998).

Because condensation domains play the predominant role in catalysis of peptide bond formation, the knowledge about its substrate tolerance is crucial for any attempt to manipulate NRPSs. Condensation domains moreover guard initiation of peptide bond formation, substrate misloading, and timing of downstream processes (Linne and Marahiel, 2000). Assays to elucidate condensation domain flexibility using an intentionally misloaded NRPS system suggested a rigid substrate recognition on acceptor site, whereas the donor amino acid is tolerated to a larger degree (Belshaw *et al.*, 1999). This observation has already been applied to construct functional hybrid NRPSs obtained by domain fusions (Doekel and Marahiel, 2000; Mootz *et al.*, 2000). Such designed NRPS systems can be used as biofactories for a defined enzymatic peptide synthesis.

Condensation domain sequences adapt to NRPS's varying domain arrangements: those that are located in the C-terminus of epimerization domains are clearly to be distinguished from those located in the C-terminus of thiolation domains. Domains involved in initiation of acyltransfer during assembly of lipopeptides represent a third group of condensation domains (Doekel and Marahiel, unpublished).

Alternatively, condensation of two intermediate carboxy acids can lead to heterocycles of the thiazoline or oxazoline type. This type of specialized condensation domain has therefore been designated as a cyclization domain (Konz *et al.*, 1997). Cyclization domains in general are found in the same domain arrangement as condensation domains. Their sufficient role in formation of heterocycles has been demonstrated *in vitro* in a yersiniabactin biosynthesis system (Gehring *et al.*, 1998b). Cyclization domains can be clearly distinguished from homologous condensation domains; interestingly, they lack their typical His motif. There are only poor data dealing with the mechanism of the cyclization reaction (Gehring *et al.*, 1998). Thiazoline structures as a product of cyclization domains can be further oxidized by the action of an optional domain that shows homologies to flavin-dependent oxyreductases. This domain type that has been identified only recently represents a 400-amino-acid-residue protein found as an insertion in adenylation domains (Molnar *et al.*, 2000) (see N-methylation domains) or in the C-terminus of thiolation domains, respectively (Silakowski *et al.*, 1999). Whereas the formation of thiazol structures therefore is also an intrinsic property of NRPSs, reduction to thiazolidine structures (see yersiniabactin) is accomplished by an external NADPH-dependent oxyreductase (Bearden *et al.*, 1997; Butterton *et al.*, 2000).

A type of domain with an epimerization activity is found as an optional insertion into minimal modules. Epimerization of carboxy acids to their D-configuration is accomplished by such epimerization domains located between a thiolation and a condensation domain (Stachelhaus and Walsh, 2000).

Another optional insertion in the domain arrangement of a minimal module was also found between adenylation and thiolation domains conferring N-methylation of substrate amino acids (Zocher and Kleinkauf, 1978). Sequence analysis of this domain type (ca. 420 amino acids) reveals similarities to other *S*-adenosylmethionine (SAM)-dependent methyltransferases (Hacker *et al.*, 2000). Obviously N-modification occurs prior to peptide bond formation (Haese *et al.*, 1993).

Only recently a domain type obviously involved in N-formylation of starter amino acids has been identified (Rouhiainen *et al.*, 2000). Therefore, this domain organization represents a new alternative in initiation of nonribosomal peptide synthesis.

Release of peptide products from modular peptide synthetases is usually accomplished by a domain found at

18069302

Metabolic Engineering 3, 64–77 (2001)
doi:10.1006/mben.2000.0170

the C-terminus of peptide synthetases. This domain type that is ca. 250 amino acids in length shows homology to thioesterases and was also found in fatty acid and polyketide synthases. They contain a GxSxG motif common for thioesterases and acyltransferases. The importance of this domain in cleavage of a peptide product could be shown by *in vivo* studies (Guenzi *et al.*, 1998b; Schneider and Marahiel, 1998). The peptide is transferred to the hydroxyl side chain of a serine in the motif GxSxG. This ester bond is subsequently cleaved by a nucleophilic attack of water (releasing a linear product) or an internal side chain of the peptide (releasing a cyclic product) (Shaw-Reid *et al.*, 1999). Recent *in vitro* studies with the recombinant isolated thioesterase domain from TycC revealed that hydrolysis via cyclization is an intrinsic factor of the thioesterase domain itself. This enzymatic domain exhibits a broad flexibility toward nonnative substrates that seems only to be restricted to the N-terminus of the peptide (Trauger *et al.*, 2000). Obviously, the selective recognition of the nucleophilic group prevents the NRPS system from the formation of mixed cyclic peptides. In addition, a proposed stereoselectivity of the ACV's thioesterase domain times and regulates the product outcome (Kallow *et al.*, 2000).

In some biosynthetic gene clusters, external thioesterase-like encoding genes were found. Gene disruption studies for external thioesterases indicated an important but not essential role in product formation (Schneider and Marahiel, 1998).

Alternatively, release of the product from the enzyme can be implemented by a reductive step yielding aldehyde or alcohol functions at the C-terminus of the peptides (Ehmann *et al.*, 1999). These reductase domains employ an NAD(P)H-dependent mechanism (Konz and Marahiel, 1999).

In summary, we have described the domain organization of peptide synthetases with a minimal module containing adenylation, thiolation, and condensation domains that bear the potential to form a single peptide bond. Minimal modules can be extended by epimerization or N-methylation domains. Alternatively, condensation domains can be replaced by cyclization domains. For release of final products, either thioesterase-like or NA(D)PH-dependent reductive domains are utilized.

DISTINCTIONS BETWEEN BACTERIAL AND FUNGAL SYSTEMS—NOVEL CHARACTERISTICS OF PEPTIDE SYNTHETASES

Among the first peptide synthetase genes that have been discovered in fungi were the ACV synthetases of *Penicillium chrysogenum* (Smith *et al.*, 1990) and the huge cyclosporin synthetase of *Tolypocladium inflatum* (Weber *et al.*, 1994). Due to the lack of introns within these large genes, it has been suggested that fungi acquired peptide synthetase genes by horizontal gene transfer from bacteria. This proposal must be revised because recently two fungal peptide synthetase genes have been discovered that bear introns: the ergotamin synthetase LPS1 of *Claviceps purpurea* (Tudzynski *et al.*, 1999) and a peptide synthetase gene of unknown function in *Metarhizium anisopliae* (Bailey *et al.*, 1996), so that peptide synthetase genes may have been present in fungi for an evolutionarily longer period than expected. In parallel, some characteristics of fungal peptide synthetases become obvious that differ from bacterial systems: most fungal peptide synthetases consist of one protein, whereas bacterial enzymes in most cases are distributed on more than one polypeptide chain (Weber *et al.*, 1994). Fungal peptide synthetases seem to share a distinct mechanism of chain termination: only in ACV synthetases is the described thioesterase domain present, whereas all other yet known fungal systems instead employ a C-terminal condensation domain (so that the number of condensation domains is $N+1$ when N is the number of peptide bonds to be formed) (Haese *et al.*, 1993). The additional C-domain is discussed to accomplish a condensation step that goes along with release of the processed peptide. Also, fungal nonribosomal synthesized peptides often contain D-amino acids, but the corresponding peptide synthetases contain no epimerization domain. In such cases, an external epimerase seems to provide a D-amino acid that is incorporated by an adenylation domain highly specific for D-configured substrates.

Some peptide synthetases like those of syringomycin (Guenzi *et al.*, 1998a) and ferrichrome comprise a domain organization that omits adenylation domains (see Table 1). Module organizations like C-T instead of C-A-T are found. It has been discussed that such additional truncated modules may be served as sites for trans-acylation reactions. In syringomycin biosynthesis, the C-terminal module is believed to be loaded with threonine that is activated on an independent one-module enzyme called SyrB. SyrB however is encoded by a gene located upstream of the 3-gene operon. If this is true, it would be the first example that breaks with the so far ubiquitar rule of a colinear assembly line of NRPS. The idea of an *in trans* acylation—a sort of cross-entry in nonribosomal peptide synthetase —could also explain the unusual domain organization found in yersiniabactin synthetase of *Yersinia pestis* (see Table 1). Here, the processed precursor peptide should contain three cysteine residues but only one adenylation domain-activating cysteine is present. Biochemical data underline that this single adenylation domain can acylate three downstream-located thiolation domains *in trans*.

Metabolic Engineering 3, 64–77 (2001)
doi:10.1006/mben.2000.0170

Domains of C-T organizations are also found in the recently characterized biosynthetic cluster for the acylheptapeptide mycosubtilin in *B. subtilis* ATCC 6633 (Duitman *et al.*, 1999) (see Table 1). In addition, the first protein of this peptide synthetase system, MycA, is a functional hybrid protein that contains parts of a fatty acid synthase and an amino acid acyltransferase. Here, MycA is believed to catalyze the synthesis of the fatty acid tail that is transferred to the first activated amino acid by the action of the amino acid transferase. It is not yet clear which role the intermediate CT domain organization plays. Nevertheless, the identification of a transferase engaged in initiating peptide synthesis of acylated peptides is new and has not been reported before. An analogous amino acid transferase coupled with a peptide synthetase has also been identified recently as part of the biosynthetic cluster of microcystin in *M. aeruginosa* (Nishizawa *et al.*, 1999).

The mycosubtilin synthetases MycB and MycC also display an unusual domain junction: whereas in intermolecular domain arrangements the condensation domain is usually found as part of the acceptor protein (an intermolecular T–C junction), in MycA and MycB the condensation domain is found as part of the donor protein providing an intermolecular C–A junction. Analogous domain arrangements have so far only been reported in the exochelin biosynthesis of *Mycobacterium smegmatis* (Fiss *et al.*, 1994).

MIXED PEPTIDE SYNTHETASES AND POLYKETIDE SYNTHASES

The striking analogy of the architecture of modular polyketide synthases and peptide synthetases has been the topic of a recently published review (Cane and Walsh, 1999). In fact, both systems represent biofactories that are combined by nature to yield mixed NRPS–PKS natural products which greatly extends the structural diversity of products.

The first identified mixed NRPS–PKS biosynthetic gene cluster was that of rapamycin of *Streptomyces hydroscopius* (Schwecke *et al.*, 1995). It contains an NRPS module for pipecolic acid incorporation into the polyketide. An analogous organization has been identified in the FK506 biosynthetic gene cluster in *Streptomyces* sp. *MA*6548 (Motamedi and Shafiee, 1998). Interestingly, in both systems two condensation domains flanking the adenylation/thiolation domain are found. The crossing of NRPS modules with up- and downstream polyketide synthase modules is guaranteed apparently by employing this domain organization.

NRPS biosynthetic gene clusters that contain polyketide synthase portions have also been identified in *Yersinia pestis*, responsible for yersiniabactin biosynthesis (Bearden *et al.*, 1997; Gehring *et al.*, 1998a). In this system, the mixed NRPS/PKS protein Irp1 contains an integral PKS module probably responsible for the incorporation of a malonyl- or acetyl-CoA extender unit linked between two thiazoline/thiazolidine moieties. It would be interesting to evaluate questions concerning how the cross-talk of these two biosystems is accomplished.

With the increasing number of identified natural product biosynthetic gene clusters, several mixed NRPS/PKS systems have recently been detected. The mycobactin NRPS of *Mycobacterium tuberculosis* (Quadri *et al.*, 1998) contains a PKS module and the microcystin biosynthesis of *M. aeruginosa* (Neilan *et al.*, 1999) utilizes a mixed NRPS/PKS machinery. The latter cluster contains a three-modular polyketide synthase that is probably involved in the biosynthesis of an uncommon amino acid based on decadienoic acid. The recently characterized biosynthetic gene clusters for epothilone and myxothiazol also employ mixed NRPS/PKS machinery (Molnar *et al.*, 2000; Silakowski *et al.*, 1999).

IMITATING NATURE'S POTENTIAL

As sequences of genes coding for peptide synthetases become readily available, detailed comparative sequence alignments were used to find an answer to questions like "what defines a module or a domain" and "what are the borders of domains?" However, a clear answer to assess domain sizes cannot be given in advance. Simple model systems to elucidate domain borders must be devised (Mootz and Marahiel, 1997; Stachelhaus *et al.*, 1996). Does a fusion of an epimerization domain to a recombinant AT-domain lead to substrate epimerization regardless of the amino acid's nature (Linne and Marahiel, unpublished)? Is any processed peptide readily cleaved by fused C-terminal thioesterase domains and, if so, how (Gokhale *et al.*, 1999a; Mootz *et al.*, 2000; Trauger *et al.*, 2000)? Can recombinant model di- or trimodular hybrid systems employing non-native condensation domains be used to catalyze peptide bond formation (Doekel and Marahiel, 2000; Mootz *et al.*, 2000)?

To foresee the potential of redesigning (un)natural products by engineered peptide synthetases, we will outline a strategy of constructing a hybrid peptide synthetase on scheme with the potential to produce a small peptide derivative. Production of a functional recombinant peptide synthetase in an approved surrogate host could overcome problems like cultivation or high-level fermentation (Kealey *et al.*, 1998). This strategy should also facilitate further

Metabolic Engineering 3, 64–77 (2001)
doi:10.1006/mben.2000.0170

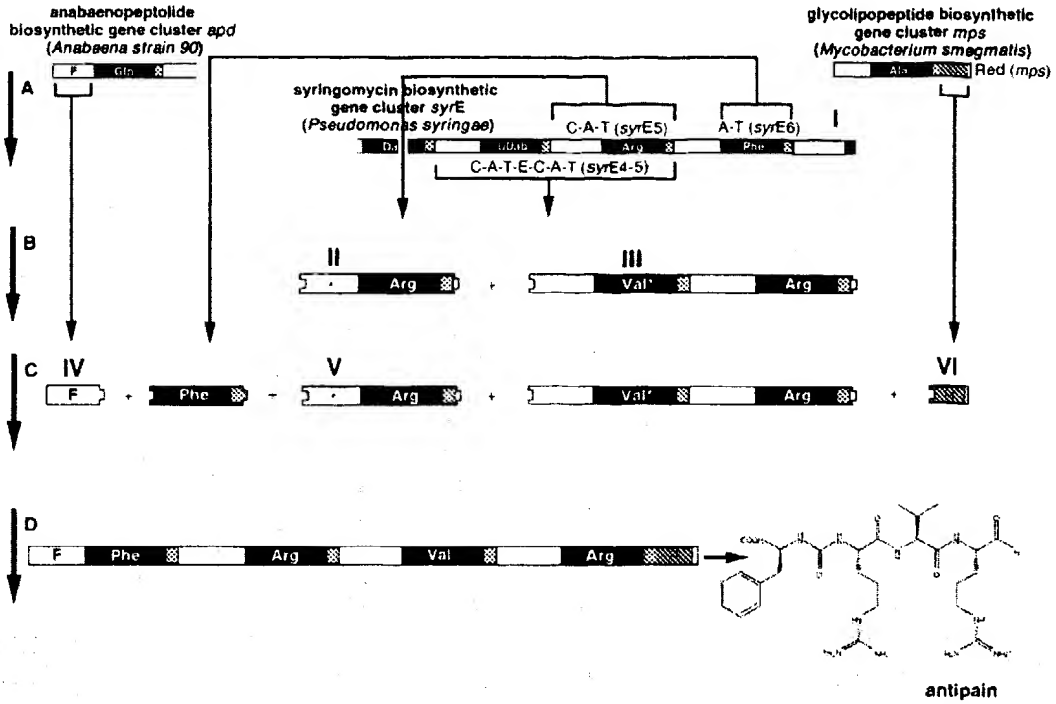


FIG. 2. Knowledge-based strategy for a targeted construction of a hybrid peptide synthetase directing the synthesis of the peptide derivative antipain. (A) The toolbox (I). Parts of the syringomycin, glycolipopeptide, and anabaenopeptolide biosynthetic gene clusters from *Pseudomonas syringae*, *Mycobacterium smegmatis*, and *Anabaena* strain 90, respectively, are illustrated. Brackets indicate regions for the virtual experiment used below. (B) Engineering steps. Site-directed mutagenesis (III) is used to alter the specificity of the D-Dab adenylation domain to Val and to make the second condensation domain condensate the carbomyl-function (II, indicated by asterisks). (C) Compiling. Gene segments encoding the domains are fused together. (D) A new template. The designed tetramodular template contains an N-terminal formylation domain implementing the N-capping of the peptide (IV). The four modules take account for the ordered incorporation and condensation of the peptide portion Phe-Arg-Val-Arg. The C-terminal reductase (indicated by a **R**) domain cleaves the processed peptide derivative from the enzyme leaving a reduced aldehyde function (VI).

genetic manipulations with convenient techniques inherent to a surrogate host.

The strategy employs operations that have already been discussed elsewhere (Mootz and Marahiel, 1999) like module and domain swapping, change-of-substrate speci ficity by mutagenesis, and an induced termination to achieve release of a defined shortened product. This strategy also involves use of a recently discovered formyl transferase (Rouhiainen *et al.*, 2000).

Antipain is a potent cathepsin inhibitor produced by cultures *of Streptomyces roseus* (Umezawa *et al.*, 1972) (see Fig. 2) whose biosynthetic machine is unknown. Its peptide portion consists of three L-amino acids with the primary sequence Arg-Val-Arg. The N-terminus is linked to the carboxy group of a N-formylated L-phenylalanine (carbamoyl function). The C-terminus is reduced to an aldehyde. It is likely that antipain is produced by a yet unknown peptide synthetase. The proposed scenario is fictive but (based on present knowledge) should generate the same compound.

The antipain peptide synthetase backbone should be tetramodular (see Fig. 2). The specificity of modules 1 to 4 should be phenylalanine, arginine, valine, and arginine. To construct such a tetramodular template, the following hybrid fusions can be employed: The modules requested can be "extracted" from the syringomycin biosynthesis operon from *Pseudomonas syringae* (I, see Fig. 2) (Guenzi *et al.*, 1998a) and subsequently fused together in a defined order to generate a hybrid template for the production of the tetrapeptide Phe-Arg-d-Dab-Arg. Functional hybrid NRPS systems have been generated recently (Doekel and Marahiel, 2000; Stachelhaus *et al.*, 1995). These artificial templates obtained by different types of domain and module fusions were shown to retain their potential to form peptides with defined structures. A rigid substrate specificity of the condensation domain toward the acceptor amino acid (Belshaw *et al.*, 1999) requires the use of the acceptor module's condensation domain [here C-A-T fusions, II, see Fig. 2 (Mootz *et al.*, 2000); T-C-A have also been shown to yield functional hybrids (Doekel and Marahiel, 2000)]. A role of interdomain linkers in maintaining communications between hybrid domains has been discussed (Gokhale *et al.*, 1999b).

Metabolic Engineering 3, 64–77 (2001)
doi:10.1006/mben.2000.0170

Several modifications must be employed whose feasibility has been confirmed generally using model NRPS systems:

(1) The specificity of the third module must be rendered from D-Dab to Val using site-directed mutagenesis, giving a template for the tetrapeptide Phe-Arg-Val-Arg (III, see Fig. 2). The elucidation of the nonribosomal code—the identification of the active-site residues involved in substrate recognition—provides a versatile tool for a predictable alteration of substrate specificity (Stachelhaus *et al.*, 1999).

(2) N-formylation in nonribosomal peptides is accomplished by a domain type that has been identified only recently (Rouhiainen *et al.*, 2000). To provide an N-formylated starter amino acid phenylalanine, the first module should be N-terminal fused to a formylation domain (IV, see Fig. 2). On the other hand, the synthesis of some N-capped nonribosomal peptides like that of the acyl peptides syringomycin (Guenzi *et al.*, 1998a) and surfactin (Cosmina *et al.*, 1993) is initiated by an N-terminal starter condensation domain.

(3) Some C-domains—for example, those found in exochelin (Fiss *et al.*, 1994; Yu *et al.*, 1998) and ACV biosynthesis (Smith *et al.*, 1990)—generate β-, δ-, or ε-peptide bonds. To aim at an antipain template, the first condensation domain used should preferably involve the carbamoyl-function for the formation of a peptide bond (V, see Fig. 2). The mechanism that governs this sort of selectivity has not yet been elucidated (Keating and Walsh, 1999).

(4) For release of the processed N-formylated tetrapeptide, a domain encoding for a NADPH-dependent reductase such as Lys2 involved in the lysine biosynthesis in yeasts (Ehmann *et al.*, 1999) and certain other peptide synthetase systems (Billman-Jacobe *et al.*, 1999) (see Table 1) should be used (VI, see Fig. 2). An induced release of such peptide products has been reported *in vivo* (de Ferra *et al.*, 1997; Guenzi *et al.*, 1998b) and—recently—*in vitro* upon fusion of thioesterase-like domains to upstream modules (Mootz *et al.*, 2000).

The proposed strategy can analogously be extended to a vast set of small peptides also to generate templates for the synthesis of vertebrate or industrial important peptides.

CONCLUSIONS

An increasing number of NRPS sequences become available. Based on these data and recent biochemical studies, a solid interpretation of the principles of NRPS's architecture and organization can now be stated. The potential of this biosynthetic machinery increases by the recent identification of integrated enzymatic domains. Studying the mechanistic principle of multidomain NRPS will consolidate attempts for their manipulation. To govern the rules of NRPS opens the way to redesign natural products and to assist nature's complexity. Based on the current knowledge, design of defined small peptides on request by combinatorially puzzled peptide synthetases becomes realistic.

ACKNOWLEDGMENTS

Work in the authors laboratory was supported by the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie. The authors are grateful to Thomas Weber for critical reading of the manuscript.

REFERENCES

Bailey, A. M., Kerchow, M. J., Hunt, B. A., Peterson, I. C., Chamley, A. K., Reynolds, S. E., and Clarkson, J. M. (1996). Cloning and sequence analysis of an intron-containing domain from a peptide synthetase-encoding gene of the entomopathogenic fungus *Metarhizium anisopliae*. *Gene* **173**, 195–197.

Bearden, S. W., Fetherston, J. D., and Perry, R. D. (1997). Genetic organization of the yersiniabactin biosynthetic region and construction of avirulent mutants in *Yersinia pestis*. *Infect. Immun.* **65**, 1659–1668.

Belshaw, P. J., Walsh, C. T., and Stachelhaus, T. (1999). Aminoacyl-CoAs as probes of condensation domain selectivity in nonribosomal peptide synthesis. *Science* **284**, 486–489.

Beyer, S., Kunze, B., Silakowski, B., and Müller, R. (1999). Metabolic diversity in myxobacteria: Identification of the myxalamid and the stigmatellin biosynthetic gene cluster of *Stigmatella aurantiaca* Sg a15 and a combined polyketide-(poly)peptide gene cluster from the epothilone producing strain *Sorangium cellulosum* So ce90. *Biochim. Biophys. Acta* **1445**, 185–195.

Billman-Jacobe, H., McConville, M. J., Haites, R. E., Kovacevic, S., and Coppel, R. L. (1999). Identification of a peptide synthetase involved in the biosynthesis of glycopeptidolipids of *Mycobacterium smegmatis*. *Mol. Microbiol.* **33**, 1244–1253.

Butterton, J. R., Choi, M. H., Watnick, P. I., Carroll, P. A., and Calderwood, S. B. (2000). *Vibrio cholerae* VibF is required for vibriobactin synthesis and is a member of the family of nonribosomal peptide synthetases. *J. Bacteriol.* **182**, 1731–1738.

Cane, D. E., and Walsh, C. T. (1999). The parallel and convergent universes of polyketide synthases and nonribosomal peptide synthetases. *Chem. Biol.* **6**, R319–R325.

Chong, P. P., Podmore, S. M., Kieser, H. M., Redenbach, M., Turgay, K., Marahiel, M., Hopwood, D. A., and Smith, C. P. (1998). Physical identification of a chromosomal locus encoding biosynthetic genes for the lipopeptide calcium-dependent antibiotic (CDA) of *Streptomyces coelicolor* A3(2). *Microbiology* **144**, 193–199.

Cole, S. T., and Barrell, B. G. (1998). Analysis of the genome of *Mycobacterium tuberculosis* H37Rv. *Novartis Found. Symp.* **217**, 160–172.

Conti, E., Stachelhaus, T., Marahiel, M. A., and Brick, P. (1997). Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. *EMBO J.* **16**, 4174–4183.

Metabolic Engineering 3, 64–77 (2001)
doi:10.1006/mben.2000.0170

Cosmina, P., Rodriguez, F., de Ferra, F., Grandi, G., Perego, M., Venema, G., and van Sinderen, D. (1993). Sequence and analysis of the genetic locus responsible for surfactin synthesis in *Bacillus subtilis*. *Mol. Microbiol.* **8**, 821–831.

de Crécy-Lagard, V., Marliere, P., and Saurin, W. (1995). Multienzymatic non ribosomal peptide biosynthesis: Identification of the functional domains catalysing peptide elongation and epimerisation. *C. R. Acad. Sci. III* **318**, 927–936.

de Crécy-Lagard, V., Saurin, W., Thibaut, D., Gil, P., Naudin, L., Crouzet, J., and Blanc, V. (1997). Streptogramin B biosynthesis in *Streptomyces pristinaespiralis* and *Streptomyces virginiae*: Molecular characterization of the last structural peptide synthetase gene. *Antimicrob. Agents Chemother.* **41**, 1904–1909.

de Ferra, F., Rodriguez, F., Tortora, O., Tosi, C., and Grandi, G. (1997). Engineering of peptide synthetases. Key role of the thioesterase-like domain for efficient production of recombinant peptides. *J. Biol. Chem.* **272**, 25304–25309.

Denessiouk, K. A., and Johnson, M. S. (2000). When fold is not important: A common structural framework for adenine and AMP binding in 12 unrelated protein families. *Proteins* **38**, 310–326.

Dittmann, E., Neilan, B. A., Erhard, M., von Döhren, H., and Borner, T. (1997). Insertional mutagenesis of a peptide synthetase gene that is responsible for hepatotoxin production in the cyanobacterium *Microcystis aeruginosa* PCC 7806. *Mol. Microbiol.* **26**, 779–787.

Doekel, S., and Marahiel, M. A. (2000). Dipeptide formation on engineered hybrid peptide synthetases. *Chem. Biol.* **7**, 373–384.

Duitman, E. H., Hamoen, L. W., Rembold, M., Venema, G., Seitz, H., Saenger, W., Bernhard, F., Reinhardt, R., Schmidt, M., Ullrich, C., Stein, T., Leenders, F., and Vater, J. (1999). The mycosubtilin synthetase of *Bacillus subtilis* ATCC6633: A multifunctional hybrid between a peptide synthetase, an amino transferase, and a fatty acid synthase. *Proc. Natl. Acad. Sci. USA* **96**, 13294–13299.

Ehmann, D. E., Gehring, A. M., and Walsh, C. T. (1999). Lysine biosynthesis in *Saccharomyces cerevisiae*: Mechanism of alpha-aminoadipate reductase (Lys2) involves posttranslational phosphopantetheinylation by Lys5. *Biochemistry* **38**, 6171–6177.

Fiechter, A. (1992). Biosurfactants: Moving toward industrial application. *Trends Biotechnol.* **10**, 208–217.

Fiss, E. H., Yu, S., and Jacobs, W. R., Jr. (1994). Identification of genes involved in the sequestration of iron in mycobacteria: The ferric exochelin biosynthetic and uptake pathways. *Mol. Microbiol.* **14**, 557–569.

Gehring, A. M., DeMoll, E., Fetherston, J. D., Mori, I., Mayhew, G. F., Blattner, F. R., Walsh, C. T., and Perry, R. D. (1998a). Iron acquisition in plague: Modular logic in enzymatic biogenesis of yersiniabactin by *Yersinia pestis*. *Chem. Biol.* **5**, 573–586.

Gehring, A. M., Mori, I., Perry, R. D., and Walsh, C. T. (1998b). The nonribosomal peptide synthetase HMWP2 forms a thiazoline ring during biogenesis of yersiniabactin, an iron-chelating virulence factor of *Yersinia pestis*. *Biochemistry* **37**, 11637–11650.

Gocht, M., and Marahiel, M. A. (1994). Analysis of core sequences in the d-Phe activating domain of the multifunctional peptide synthetase TycA by site-directed mutagenesis. *J. Bacteriol.* **176**, 2654–2662.

Gokhale, R. S., Hunziker, D., Cane, D. E., and Khosla, C. (1999a). Mechanism and specificity of the terminal thioesterase domain from the erythromycin polyketide synthase. *Chem. Biol.* **6**, 117–125.

Gokhale, R. S., Tsuji, S. Y., Cane, D. E., and Khosla, C. (1999b). Dissecting and exploiting intermodular communication in polyketide synthases. *Science* **284**, 482–485.

Guenzi, E., Galli, G., Grgurina, I., Gross, D. C., and Grandi, G. (1998a). Characterization of the syringomycin synthetase gene cluster. A link between prokaryotic and eukaryotic peptide synthetases. *J. Biol. Chem.* **273**, 32857–32863.

Guenzi, E., Galli, G., Grgurina, I., Pace, E., Ferranti, P., and Grandi, G. (1998b). Coordinate transcription and physical linkage of domains in surfactin synthetase are not essential for proper assembly and activity of the multienzyme complex. *J. Biol. Chem.* **273**, 14403–14410.

Gutierrez, S., Diez, B., Montenegro, E., and Martin, J. F. (1991). Characterization of the *Cephalosporium acremonium* pcbAB gene encoding alpha-aminoadipyl-cysteinyl-valine synthetase, a large multidomain peptide synthetase: Linkage to the pcbC gene as a cluster of early cephalosporin biosynthetic genes and evidence of multiple functional domains. *J. Bacteriol.* **173**, 2354–2365.

Hacker, C., Glinski, M., Hornbogen, T., Doller, A., and Zocher, R. (2000). Mutational analysis of the N-methyltransferase domain of the multifunctional enzyme enniatin synthetase. *J. Biol. Chem.*, in press.

Haese, A., Pieper, R., von Ostrowski, T., and Zocher, R. (1994). Bacterial expression of catalytically active fragments of the multifunctional enzyme enniatin synthetase. *J. Mol. Biol.* **243**, 116–122.

Haese, A., Schubert, M., Herrmann, M., and Zocher, R. (1993). Molecular characterization of the enniatin synthetase gene encoding a multifunctional enzyme catalysing N-methyldepsipeptide formation in *Fusarium scirpi*. *Mol. Microbiol.* **7**, 905–914.

Higashijima, T., Inubushi, T., Ueno, T., and Miyazawa, T. (1979). NMR saturation transfer and line shape analyses of cyclic tetradepsipeptide AM toxin II: Conformational equilibrium with very unequal populations. *FEBS Lett.* **105**, 337–340.

Johnson, R. D., Johnson, L., Itoh, Y., Kodama, M., Otani, H., and Kohmoto, K. (2000). Cloning and characterization of a cyclic peptide synthetase gene from *Alternaria alternata* apple pathotype whose product is involved in AM-toxin synthesis and pathogenicity. *Mol. Plant Microbe Interact.* **13**, 742–753.

Julien, B., Shah, S., Ziermann, R., Goldman, R., Katz, L., and Khosla, C. (2000). Isolation and characterization of the epothilone biosynthetic gene cluster from *Sorangium cellulosum*. *Gene* **249**, 153–160.

Kallow, W., Kennedy, J., Arezi, B., Turner, G., and von Döhren, H. (2000). Thioesterase domain of delta-(l-alpha-aminoadipyl)-*l*-cysteinyl-*d*-valine synthetase: Alteration of stereospecificity by site-directed mutagenesis. *J. Mol. Biol.* **297**, 395–408.

Kealey, J. T., Liu, L., Santi, D. V., Betlach, M. C., and Barr, P. J. (1998). Production of a polyketide natural product in nonpolyketide-producing prokaryotic and eukaryotic hosts. *Proc. Natl. Acad. Sci. USA* **95**, 505–509.

Keating, T. A., and Walsh, C. T. (1999). Initiation, elongation, and termination strategies in polyketide and polypeptide antibiotic biosynthesis. *Curr. Opin. Chem. Biol.* **3**, 598–606.

Konz, D., Klens, A., Schörgendorfer, K., and Marahiel, M. A. (1997). The bacitracin biosynthesis operon of *Bacillus licheniformis* ATCC 10716: Molecular characterization of three multi-modular peptide synthetases. *Chem. Biol.* **4**, 927–937.

Konz, D., and Marahiel, M. A. (1999). How do peptide synthetases generate structural diversity? *Chem. Biol.* **6**, R39–R48.

Lambalot, R. H., Gehring, A. M., Flugel, R. S., Zuber, P., LaCelle, M., Marahiel, M. A., Reid, R., Khosla, C., and Walsh, C. T. (1996). A new enzyme superfamily—The phosphopantetheinyl transferases. *Chem. Biol.* **3**, 923–936.

Linne, U., and Marahiel, M. A. (2000). Control of directionality in nonribosomal peptide synthesis. The role of condensation domain in preventing mis-initiation and timing of epimerization. *Biochemistry* **39**(34), 10439–10447.

Marahiel, M. A., Stachelhaus, T., and Mootz, H. D. (1997). Modular peptide synthetases involved in non-ribosomal peptide synthesis. *Chem. Rev.* **97**, 2651–2673.

Metabolic Engineering 3, 64–77 (2001)
doi:10.1006/mben.2000.0170

Mei, B., Budde, A. D., and Leong, S. A. (1993). sid1, a gene initiating siderophore biosynthesis in *Ustilago maydis*: Molecular characterization, regulation by iron, and role in phytopathogenicity. *Proc. Natl. Acad. Sci. USA* **90**, 903–907.

Molnar, I., Schupp, T., Ono, M., Zirkle, R. E., Milnamow, M., Nowak-Thompson, B., Engel, N., Toupet, C., Stratmann, A., Cyr, D. D., Gorlach, J., Mayo, J. M., Hu, A., Goff, S., Schmid, J., and Ligon, J. M. (2000). The biosynthetic gene cluster for the microtubule-stabilizing agents epothilones A and B from *Sorangium cellulosum* So CE90. *Chem. Biol.* **7**, 97–109.

Mootz, H. D., and Marahiel, M. A. (1997). The tyrocidine biosynthesis operon of *Bacillus brevis*: Complete nucleotide sequence and biochemical characterization of functional internal adenylation domains. *J. Bacteriol.* **179**, 6843–6850.

Mootz, H. D., and Marahiel, M. A. (1999). Design and application of multimodular peptide synthetases. *Curr. Opin. Biotechnol.* **10**, 341–348.

Mootz, H. D., Schwarzer, D., and Marahiel, M. A. (2000). Construction of hybrid peptide synthetases by module and domain fusions. *Proc. Natl. Acad. Sci. USA* **97**, 5848–5853.

Motamedi, H., and Shafiee, A. (1998). The biosynthetic gene cluster for the macrolactone ring of the immunosuppressant FK506. *Eur. J. Biochem.* **256**, 528–534.

Nakano, M. M., Corbell, N., Besson, J., and Zuber, P. (1992). Isolation and characterization of sfp: A gene that functions in the production of the lipopeptide biosurfactant, surfactin, in *Bacillus subtilis*. *Mol. Gen. Genet.* **232**, 313–321.

Neilan, B. A., Dittmann, E., Rouhiainen, L., Bass, R. A., Schaub, V., Sivonen, K., and Borner, T. (1999). Nonribosomal peptide synthesis and toxigenicity of cyanobacteria. *J. Bacteriol.* **181**, 4089–4097.

Neilan, B. A., Jacobs, D., Del Dot, T., Blackall, L. L., Hawkins, P. R., Cox, P. T., and Goodman, A. E. (1997). rRNA sequences and evolutionary relationships among toxic and nontoxic cyanobacteria of the genus Microcystis. *Int. J. Syst. Bacteriol.* **47**, 693–697.

Nishizawa, T., Asayama, M., Fujii, K., Harada, K., and Shirai, M. (1999). Genetic analysis of the peptide synthetase genes for a cyclic heptapeptide microcystin in microcystis spp. *J. Biochem. (Tokyo)* **126**, 520–529.

Pavela-Vrancic, M., Pfeifer, E., Schröder, W., von Döhren, H., and Kleinkauf, H. (1994). Identification of the ATP binding site in tyrocidine synthetase I by selective modification with fluorescein 5′-isothiocyanate. *J. Biol. Chem.* **269**, 14962–14966.

Quadri, L. E., Sello, J., Keating, T. A., Weinreb, P. H., and Walsh, C. T. (1998). Identification of a *Mycobacterium tuberculosis* gene cluster encoding the biosynthetic enzymes for assembly of the virulence-conferring siderophore mycobactin. *Chem. Biol.* **5**, 631–645.

Redenbach, M., Kieser, H. M., Denapaite, D., Eichner, A., Cullum, J., Kinashi, H., and Hopwood, D. A. (1996). A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb *Streptomyces coelicolor* A3(2) chromosome. *Mol. Microbiol.* **21**, 77–96.

Reuter, K., Mofid, M. R., Marahiel, M. A., and Ficner, R. (1999). Crystal structure of the surfactin synthetase-activating enzyme sfp: A prototype of the 4′-phosphopantetheinyl transferase superfamily. *EMBO J.* **18**, 6823–6831.

Riederer, B., Han, M., and Keller, U. (1996). d-Lysergyl peptide synthetase from the ergot fungus *Claviceps purpurea*. *J. Biol. Chem.* **271**, 27524–27530.

Rouhiainen, L. L. P., Suomalainen, S., Hyytiäinen, H., Buikema, W. J., Hasselkorn, R., and Sivonen, K. (2000). Genes encoding synthetases of cyclic depsipeptides, Anabaenopeptilides, in *Anabaena* strain 90. *Mol. Microbiol.* **37**(1), 156–167.

Schneider, A., and Marahiel, M. A. (1998). Genetic evidence for a role of thioesterase domains, integrated in or associated with peptide synthetases, in non-ribosomal peptide biosynthesis in *Bacillus subtilis*. *Arch. Microbiol.* **169**, 404–410.

Schwecke, T., Aharonowitz, Y., Palissa, H., von Döhren, H., Kleinkauf, H., and van Liempt, H. (1992). Enzymatic characterisation of the multifunctional enzyme delta-(l-alpha-aminoadipyl)-l-cysteinyl-d-valine synthetase from *Streptomyces clavuligerus*. *Eur. J. Biochem.* **205**, 687–694.

Schwecke, T., Aparicio, J. F., Molnar, I., König, A., Khaw, L. E., Haydock, S. F., Oliynyk, M., Caffrey, P., Cortes, J., Lester, J. B., *et al.* (1995). The biosynthetic gene cluster for the polyketide immunosuppressant rapamycin. *Proc. Natl. Acad. Sci. USA* **92**, 7839–7843.

Scott-Craig, J. S., Panaccione, D. G., Pocard, J. A., and Walton, J. D. (1992). The cyclic peptide synthetase catalyzing HC-toxin production in the filamentous fungus *Cochliobolus carbonum* is encoded by a 15.7-kilobase open reading frame. *J. Biol. Chem.* **267**, 26044–26049.

Shaw-Reid, C. A., Kelleher, N. L., Losey, H. C., Gehring, A. M., Berg, C., and Walsh, C. T. (1999). Assembly line enzymology by multimodular nonribosomal peptide synthetases: The thioesterase domain of *E. coli* EntF catalyzes both elongation and cyclolactonization. *Chem. Biol.* **6**, 385–400.

Silakowski, B., Schairer, H. U., Ehret, H., Kunze, B., Weinig, S., Nordsiek, G., Brandt, P., Blocker, H., Hofle, G., Beyer, S., and Müller, R. (1999). New lessons for combinatorial biosynthesis from myxobacteria. The myxothiazol biosynthetic gene cluster of *Stigmatella aurantiaca* dw4/3-1. *J. Biol. Chem.* **274**, 37391–37399.

Smith, D. J., Earl, A. J., and Turner, G. (1990). The multifunctional peptide synthetase performing the first step of penicillin biosynthesis in *Penicillium chrysogenum* is a 421,073 dalton protein similar to *Bacillus brevis* peptide antibiotic synthetases. *EMBO J.* **9**, 2743–2750.

Stachelhaus, T., Hüser, A., and Marahiel, M. A. (1996). Biochemical characterization of peptidyl carrier protein (PCP), the thiolation domain of multifunctional peptide synthetases. *Chem. Biol.* **3**, 913–921.

Stachelhaus, T., Mootz, H. D., Bergendahl, V., and Marahiel, M. A. (1998). Peptide bond formation in nonribosomal peptide biosynthesis. Catalytic role of the condensation domain. *J. Biol. Chem.* **273**, 22773–22781.

Stachelhaus, T., Mootz, H. D., and Marahiel, M. A. (1999). The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem. Biol.* **6**, 493–505.

Stachelhaus, T., Schneider, A., and Marahiel, M. A. (1995). Rational design of peptide antibiotics by targeted replacement of bacterial and fungal domains. *Science* **269**, 69–72.

Stachelhaus, T., and Walsh, C. T. (2000). Mutational analysis of the epimerization domain in the initiation module PheATE of gramicidin S synthetase. *Biochemistry* **39**, 5775–5787.

Stein, T., Vater, J., Kruft, V., Otto, A., Wittmann-Liebold, B., Franke, P., Panico, M., McDowell, R., and Morris, H. R. (1996). The multiple carrier model of nonribosomal peptide biosynthesis at modular multienzymatic templates. *J. Biol. Chem.* **271**, 15428–15435.

Stindl, A., and Keller, U. (1994). Epimerization of the d-valine portion in the biosynthesis of actinomycin D. *Biochemistry* **33**, 9358–9364.

Suvarna, K., Seah, L., Bhattacherjee, V., and Bhattacharjee, J. K. (1998). Molecular analysis of the LYS2 gene of *Candida albicans*: Homology to peptide antibiotic synthetases and the regulation of the alpha-aminoadipate reductase. *Curr. Genet.* **33**, 268–275.

Tang, L., Shah, S., Chung, L., Carney, J., Katz, L., Khosla, C., and Julien, B. (2000). Cloning and heterologous expression of the epothilone gene cluster. *Science* **287**, 640–642.

Metabolic Engineering 3, 64–77 (2001)
doi:10.1006/mben.2000.0170

Trauger, J. W., Kohli, R. M., Mootz, H. D., Marahiel, M. A., and Walsh, C. T. (2000). Peptide cyclization catalyzed by the thioesterase domain of tyrocidine synthetase. *Nature* **407**, 215–218.

Trauger, J. W., and Walsh, C. T. (2000). Heterologous expression in *Escherichia coli* of the first module of the nonribosomal peptide synthetase for chloroeremomycin, a vancomycin-type glycopeptide antibiotic. *Proc. Natl. Acad. Sci. USA* **97**, 3112–3117.

Tudzynski, P., Holter, K., Correia, T., Arntz, C., Grammel, N., and Keller, U. (1999). Evidence for an ergot alkaloid gene cluster in *Claviceps purpurea*. *Mol. Gen. Genet.* **261**, 133–141.

Turgay, K., and Marahiel, M. A. (1994). A general approach for identifying and cloning peptide synthetase genes. *Pept. Res.* **7**, 238–241.

Umezawa, S., Tatsuta, K., Fujimoto, K., Tsuchiya, T., and Umezawa, H. (1972). Structure of antipain, a new Sakaguchi-positive product of Streptomyces. *J. Antibiot. (Tokyo)* **25**, 267–270.

van Wageningen, A. M., Kirkpatrick, P. N., Williams, D. H., Harris, B. R., Kershaw, J. K., Lennard, N. J., Jones, M., Jones, S. J., and Solenberg, P. J. (1998). Sequencing and analysis of genes involved in the biosynthesis of a vancomycin group antibiotic. *Chem. Biol.* **5**, 155–162.

Weber, G., Schörgendorfer, K., Schneider-Scherzer, E., and Leitner, E. (1994). The peptide synthetase catalyzing cyclosporine production in *Tolypocladium niveum* is encoded by a giant 45.8-kilobase open reading frame. *Curr. Genet.* **26**, 120–125.

Weber, T., Baumgartner, R., Renner, C., Marahiel, M. A., and Holak, T. A. (2000). Solution structure of PCP, a prototype for the peptidyl carrier domains of modular peptide synthetases. *Struct. Fold Des.* **8**, 407–418.

Weinreb, P. H., Quadri, L. E., Walsh, C. T., and Zuber, P. (1998). Stoichiometry and specificity of *in vitro* phosphopantetheinylation and aminoacylation of the valine-activating module of surfactin synthetase. *Biochemistry* **37**, 1575–1584.

Wyckoff, E. E., Stoebner, J. A., Reed, K. E., and Payne, S. M. (1997). Cloning of a *Vibrio cholerae* vibriobactin gene cluster: Identification of genes required for early steps in siderophore biosynthesis. *J. Bacteriol.* **179**, 7055–7062.

Yu, S., Fiss, E., and Jacobs, W. R., Jr. (1998). Analysis of the exochelin locus in *Mycobacterium smegmatis*: Biosynthesis genes have homology with genes of the peptide synthetase family. *J. Bacteriol.* **180**, 4676–4685.

Zhu, W., Arceneaux, J. E., Beggs, M. L., Byers, B. R., Eisenach, K. D., and Lundrigan, M. D. (1998). Exochelin genes in *Mycobacterium smegmatis*: Identification of an ABC transporter and two nonribosomal peptide synthetase genes. *Mol. Microbiol.* **29**, 629–639.

Zocher, R., and Kleinkauf, H. (1978). Biosynthesis of enniatin B: Partial purification and characterization of the synthesizing enzyme and studies of the biosynthesis. *Biochem. Biophys. Res. Commun.* **81**, 1162–1167.